



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Run on: Sat May 13 08:34:56 2000; MasPar time 21.20 Seconds
659.132 Million cell updates/sec
Tabular output not generated.

Title: >US-09-331-631-8
Description: (1-590) from US09331631.pep (1 of 4)
Perfect Score: 4332
Sequence: 1 MVRNKSACVLLFSLFLSFC.....FNSNPQESFYVSRQRQRASE 590

Scoring table: PAM 150
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:geneseqp

Statistics: Mean 36.568; Variance 172.610; scale 0.212

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	4332	100.0	590	1	Gossypium hirsutum ant	0.00e+00
2	2101	48.5	566	1	Sequence encoded by 67	2.30e-174
3	1910	44.1	525	1	Theobroma cacao anti	4.54e-157
4	1246	28.8	666	1	Macadamia integrifolia	3.18e-97
5	1244	28.7	666	1	Macadamia integrifolia	4.80e-97
6	1236	28.5	666	1	Macadamia integrifolia	2.50e-96
7	966	22.3	605	1	Glycine max antimicrob	3.32e-72
8	955	22.0	593	1	zea mays antimicrobia	3.17e-71
9	834	19.3	614	1	peanut allergen Ara hi	1.73e-60
10	834	19.3	614	1	Arachis hypogaea anti	1.73e-60
11	784	18.1	626	1	Peanut allergen Ara hi	4.58e-56
12	761	17.6	409	1	G. max truncated SBP2	4.88e-54
13	764	17.6	489	1	G. max SBP2 protein.	2.65e-54
14	758	17.5	637	1	Hordeum vulgare anti	8.97e-54
15	737	17.0	524	1	G. max SBP1 protein.	6.35e-52
16	716	16.5	444	1	G. max truncated SBP1	4.48e-50
17	190	4.4	395	1	Mouse SRY-related prot	1.65e-02
18	147	3.4	1382	1	Human metastasis-asso	2.79e-02
19	143	3.3	351	1	Sequence of Histidine	5.45e-02
20	142	3.3	515	1	Oat globin A2B.	6.43e-02
21	138	3.2	186	1	Trypanosoma cruzi anti	1.25e-01
22	136	3.1	86	1	GST-HD fusion protein	1.73e-01
23	136	3.1	86	1	GST-HD fusion protein	1.73e-01

24	136	3.1	94	1	W95080	GST-HD fusion protein	1.73e-01
25	136	3.1	94	1	W95075	GST-HD fusion protein	1.73e-01
26	134	3.1	561	1	P61363	Soybean glycinin A5A4B	2.41e-01
27	133	3.1	562	1	R70491	Leucocytosom protozoa	2.83e-01
28	132	3.0	472	1	R7127	Modified oat globin A2	3.33e-01
29	125	2.9	499	1	P82755	Rice storage protein.	1.03e+00
30	126	2.9	1529	1	R97985	CORC potassium channel	8.80e-01
31	124	2.9	1931	1	R27649	Human calcium channel	1.21e+00
32	120	2.8	96	1	W99832	HIV A305 protein seque	2.29e+00
33	123	2.8	108	1	W95071	Amino acid sequence of	1.42e+00
34	123	2.8	108	1	W95076	Amino acid sequence of	1.42e+00
35	121	2.8	167	1	W89768	Staphylococcus aureus	1.96e+00
36	123	2.8	516	1	P61362	Soybean glycinin A3B4	1.42e+00
37	123	2.8	737	1	W18317	Drosophila Deltex prot	1.42e+00
38	123	2.8	737	1	R76639	Deltex protein produce	1.42e+00
39	123	2.8	737	1	R76640	Deltex protein.	1.42e+00
40	122	2.8	1132	1	R97866	Chicken Leucocytosom	1.67e+00
41	120	2.8	1343	1	W18666	Mouse metastasis-asso	2.29e+00
42	119	2.7	434	1	R96420	Peptide fragment of N-	2.69e+00
43	119	2.7	2237	1	W63142	Human calcium channel	2.69e+00
44	119	2.7	2237	1	R71006	Human neuronal calcium	2.69e+00
45	119	2.7	2339	1	R33549	Sequence of the alpha	2.69e+00

ALIGNMENTS

RESULT	1	ALIGNMENTS
ID	W62832; standard; Protein: 590 AA.	
AC	W62832;	
DT	27-OCT-1998 (first entry)	
DE	Gossypium hirsutum antimicrobial protein.	
KW	antimicrobial protein; infestation; control.	
OS	Gossypium hirsutum.	
PN	W09827805-A1.	
PD	02-JUL-1998.	
PF	22-DEC-1997; AU0874.	
PR	20-DEC-1996; AU-004275.	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
PI	Bower NT, Goulter KC, Green JL, Mannes JM, Marcus JP;	
DR	WPI: 98-377279/32.	
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -	
PT	useful for controlling microbial infestations of plants or mammals	
PS	Claim 1; Page 49-51; 96pp; English.	
CC	The sequence is that of an antimicrobial protein which can	
CC	be used to control microbial infestations in plants and mammalian	
CC	animals.	
Sequence	590 AA:	
Query Match	100.0%	Score 4332; DB 1; Length 590;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;
Matches	590; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DB	1 MVRNKSACVLLFSLFLSFCALCSAKDPFRGDDDPFRYEDCRRRCWDRGQEQOQ 60	
DB	1 MVRNKSACVLLFSLFLSFCALCSAKDPFRGDDDPFRYEDCRRRCWDRGQEQOQ 60	
QY	1 MVRNKSACVLLFSLFLSFCALCSAKDPFRGDDDPFRYEDCRRRCWDRGQEQOQ 60	
DB	61 CESKCSQYGEKDDQORHNPEDPORRYECQECRCQEQEQPOCCQRLKRFEEQOQS 120	
QY	61 CESKCSQYGEKDDQORHNPEDPORRYECQECRCQEQEQPOCCQRLKRFEEQOQS 120	
DB	121 OROFOECQOHCQOERPEKQOCVRECKRYOENPWREREEAEETEGEGEQOSH 180	
QY	121 OROFOECQOHCQOERPEKQOCVRECKRYOENPWREREEAEETEGEGEQOSH 180	
DB	181 PHFHRFSQSRFRREHGNFVLQRPASRNPILRGINERLSILEANPTFVLPHHCDAE 240	
QY	181 PHFHRFSQSRFRREHGNFVLQRPASRNPILRGINERLSILEANPTFVLPHHCDAE 240	
DB	241 KIYLVNNGRGLTFLTHENKESYNIYGVVYKVPAGSTYLLANONKKEILIAVLRPN 300	
QY	241 KIYLVNNGRGLTFLTHENKESYNIYGVVYKVPAGSTYLLANONKKEILIAVLRPN 300	
DB	301 NPGQEEFFPAGSORPQSYLRAFSREILLEPAFNTRESEUDELFGROSRROOGGMPFK 360	

Qy	301	NPGE	EEFFPAGSQRPQSLYAFNR	ELIEPAFNTR	RSQDLDELFGCR	SRRRQGGGMRK	360	
Db	361	ASQD	IRALSGEARS	PRKSGERRAFNL	STPRYSNONG	RFEEACPEEROLDINVTY	420	
Qy	361	ASQD	IRALSGEARS	PRKSGERRAFNL	STPRYSNONG	RFEEACPEEROLDINVTY	420	
Db	421	SALD	NOGSIIVPHYSKATP	VIIVLTGNGAYE	VSPLPROSSIEE	EEDEEEDDEE	480	
Qy	421	SALD	NOGSIIVPHYSKATP	VIIVLTGNGAYE	VSPLPROSSIEE	EEDEEEDDEE	480	
Db	481	EERR	SGOYKTR	RSRLSRD	IVVPANP	PYFVASONONL	PLRTGGGLGNQININDHNRIF	540
Qy	481	EERR	SGOYKTR	RSRLSRD	IVVPANP	PYFVASONONL	PLRTGGGLGNQININDHNRIF	540
Db	541	VAGK	INHHYRQWDSQAK	ELAFVSSRLV	DELFINSPQ	STFYVSRO	ROPA	590
Qy	541	VAGK	INHHYRQWDSQAK	ELAFVSSRLV	DELFINSPQ	STFYVSRO	ROPA	590

RESULT 2
ID R20181 standard; Protein; 566 AA.
AC R20181;
DT 16-APR-1992 (first entry)
DE Sequence encoded by 67 kd T. cacao protein cDNA.
KW Cocoa, flavour; vicilin; seed storage protein.
OS Theobroma cacao.
PN WO9119801-A.
PD 26-DEC-1991.
PE 07-JUN-1991; GB-013016.
PR 11-JUN-1990; GB-013016.
PA (MRS C) MARS UK LTD.
PI Spencer ME, Hodge R, Deakin EA, Ashton S;
DR WPI: 92-024418/03.
N-PSDB: 020337.
PT Recombinant cocoa proteins - are responsible for flavour in cocoa
beans and produced in large quantities using yeast and bacterial
PT expression vectors
PT Claim 4; Fig 2; 59pp; English.
CC The inventors claim a 67 kd and 31 kd T. cacao protein, and
fragments, and encoding DNAs. The 47 kd and 31 kd proteins are
detected from the 67 kd precursor. T. cacao protein cDNA was
detected in a cDNA library prepared from immature cocoa beans RNA
CC using a probe based on the AA sequence of a CNBR peptide common to
the 47 kd and 31 kd polypeptides. Homology searches revealed close
homologies between the 67 kd polypeptide and the vicilins, which are
CC seed storage proteins.
CC Sequence 566 AA;
Q

Only Match	48.5%	Score 2101	DB 1	Length 566																																																		
Best Local Similarity	57.8%	Pred. No. 2.30e-174																																																				
Matches 301	Conservative 104	Mismatches 97	Indels 19	Gaps 13																																																		
D	D	P	R	O	O	O	E	C	C	R	C	E	A	T	E	R	E	O	D	O	C	E	R	E	K	E	O	R	O	O	E	E	L	O	R	O	O	C	R	C	O	E	O	96										
Q	Y	82	D	P	O	R	O	R	E	E	C	O	D	E	C	-	R	O	-	E	E	R	O	O	C	O	C	R	L	K	R	-	E	O	-	O	O	-	O	R	O	O	H	C	O	135								
D	97	O	G	O	R	E	O	O	O	C	O	R	K	C	M	E	O	R	K	E	O	-	R	G	E	H	E	N	Y	N	H	K	N	R	S	E	E	E	G	O	O	R	N	N	P	P	R	K	S	O	T	R	155	
Q	136	Q	R	P	E	K	K	O	O	C	A	V	E	C	E	K	E	O	E	N	M	R	E	R	E	-	E	A	E	E	E	-	T	E	G	E	O	E	S	N	P	F	H	N	-	R	S	O	S	R	192			
D	156	F	R	D	E	B	G	N	K	I	O	R	P	A	E	N	S	P	L	K	G	I	N	D	R	L	A	M	E	A	N	P	T	F	I	P	H	H	C	A	E	A	I	E	V	T	N	G	K	T	215			
Q	193	F	R	E	H	G	N	F	R	I	O	R	E	A	S	R	P	I	D	R	G	I	N	E	F	L	S	T	I	L	E	A	N	P	T	F	I	P	H	C	A	E	I	L	V	T	N	G	K	T	252			
D	216	T	E	F	V	H	E	N	E	S	T	I	N	V	O	G	F	V	S	P	A	S	T	A	S	T	I	V	A	S	O	D	N	O	K	E	I	T	I	V	L	P	N	S	P	G	K	E	V	P	P	A	G	275
Q	253	T	E	F	L	H	E	K	E	S	T	I	N	V	O	G	F	V	S	P	A	S	T	I	V	L	A	S	O	D	N	O	K	E	I	T	I	V	L	P	N	S	P	G	K	E	V	P	P	A	G	312		
D	276	N	K	E	E	S	T	I	G	A	F	S	T	E	V	L	E	N	T	O	R	E	K	E	I	E	E	B	O	G	O	R	O	O	G	O	G	M	F	R	K	A	P	E	O	I	R	A	S	O	335			
Y	313	S	O	R	O	S	T	R	A	R	S	R	I	L	E	A	F	A	T	R	E	O	L	D	E	L	G	G	S	O	R	R	O	O	-	G	M	R	K	S	O	E	D	I	R	A	S	O	371					

[illegible]

TD RESULT 3
 TD W62831 standard; Protein; 525 AA.
 AC W62831;
 DT 27-Oct-1998 (first entry)
 DE Theobroma cacao antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Theobroma cacao.
 PN M09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; A00874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI; 98-377279/32.
 PT Novel anti-microbial protein from e.g. *Macadamia integrifolia* -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 47-49; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 CC be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 525 AA;

Query Match	44.1%	Score 1910	DB 1	Length 525
Best Local Similarity	56.9%	Pred. No. 4,54e-157		
Matches 261	Conservative	95	Mismatches 92	Indels 11
			Gaps 10	
Db	DPRQDYEQCCRCSEASEATEREERQCEQRCERERKCYEQROOQROOEELQROYOQOCQRCQEQ	96		
Qy	82 DPQRREYECQCEQEC-RQO-EEHQRPQCQQRCLKRR-BDE-QQOS-QRQFOGCOHCHQOE	135		
Dp	97 QGQREQOQCCQRCCKWEQYQKEDE-RGEHENYHNHKKNSDEEBEGQORNNPYTFPKRRSFQTR	155		
Qy	136 QRPEKQOQCVECEKEKQOEPMREEREB-EAEDEE-TEEGEOESHPHFHH-RRRSQSR	192		
Dp	156 FRDEFGNCKILQREPAENSPLPKGINDRIAMFEANPMTFTLPHHCDAEALFYVNGKTI	215		
Qy	193 FREEHGNFRVLQRRASRHPILRGINERLSTILENPNPTFVLPHHCDCEKILYLVINGGTL	252		
Dp	216 TPTVTHENKESINVQRTIVSVSPAGSTVYVVSQDNQEKLTAVIALPVNSPKYELFFPAG	275		
Qy	253 TELTHENKESNINIVPGVYVVKPAPASTYVLANDQNEKLTIAVLHPRVNNNGQEEFFEPAG	312		
Dp	276 NKKPESTYGAATSVYLEVFTVPTQOEXLELLEBQRGOKRQOOGQGMRRKAKPEQIRASIQ	335		
Qy	313 SQRPQSTIRARSRLELLEPAFTNRSEQIDLELFGHQSRRRQOQ-GMRKKSQOEQIRLASQ	371		
Dp	336 QATSPRRHGRERLAINLSSQSPVYSNONGRFEACPEDFSOFQMDVYASAKLNGAIF	395		
Qy	372 EATSPREKSGERFAPNLISQTPRYSNONGRFEACPEEPQLORDINTVYALQNGSIF	431		
Dp	396 VPHNNSKATFVFTYTDGQGYAQMACPHLSSQSQSQSQSGRODRRQOEESDEETFGQYQY	455		
Qy	432 VPHNNSKATFVILVTEGNGYAEWSPMLPROSSYEEDDEEBEQOEDEERSGQYRKTI	491		
Dp	456 KAPLSPGQVFEVAPAGHAVTFEASQDKQPLNVAEFLNQN 494			

Query	567	VDEIFNSNPQESFYVS-RQRRAS	589
Db	624	VEELFNSODESIFPPGPRHOQOOS	647
Oy	508	PVTIVASQNLRLMTGGLIYVNTNPNPDHNOITVAGKINHV-RQMSQAKELAFGVSSRL	566
Oy	569	PVVFVSSGENKULLFAEGINQON-N-H-EN-FLAGERANVLQIEEQANIELFAAPRKE	623
Db	448	GNQVAEWVSPILPROSSVEEEDDEEEOQEEERSSQYRKIRSLTSGDIFVVPANF	507
Oy	517	GEADVACAPHL---SG-RHGGRGGRK-HEEED--V-HYEOVKAPLSREALIVPVGH	568
Db	388	LLSQPRRNSQNGRFFECAPREFPQRLRINTVTSALQNGSIVPVHYNKATFVILYTE	447
Oy	274	PAGSVIVYLANODNKEKLIIVLHVRPVNPGQFEFFPAGSORPSYLRASRELLIEPAFN	333
Db	404	TOAEKLRGVLG--Q--QRE---GYIIASQOIIELRDSESRMRHIRGGESESRQPN	456
Oy	334	TRSELDLDELFEGRSRRRQOGQMFRRASQEQIALISQEAET-SPR-E--KSGE--RPAFN	387
Db	457	LEFKRPLISNKGAYEKKPEDYQLODMDSVFIANITQSGNMKGPFENRSTKVVVAS	516
Oy	214	RGINEFRSLILEANPNFTVLPDHODAETIYLVYNGRGLTFLETHENKESINIVPGVYKV	273
Db	284	RALKRYLVLLLEAPNNAFVLPTHLDADAAILLVYTGSRALKIHHNDRESYNLECGDVIRI	343
Oy	156	PWRGE-REEEAEEET--EEGQEQSHNPFHFRSFSQREERHGNPRVLORFASRPIL	213
Db	224	GRGGLINPQNGSGRYDEGEKESQNDPIYDEDSLSLTPRTTEGCHLSYLENTYGRSKLL	283
Oy	96	QOEERQPOQCORCRKRFPEEOOOSOFORFCQCHQEOERPEKKQCYVRECREKYQEN	155
Db	171	EDEEK---YEER-MKEEDNRKRDQOAREYEDRCRRCEQOERPOQ--YOCORRCQEOQRH	223
Oy	36	DPKPRTEDCRRRCWMDTRGQKEQOQCEESCSQYGEKQDQQRHNEPDRRECEQOECR	95
Db	121	DPQOYEQCQRCQRHREPRRHMTCCQRCGRRY-EKEK--R-R-Q--QKRYEP-QQ--R	170
Oy	222	Conservative 151; Mismatches 144; Indels 47; Caps 30;	
Query Match	28.5%;	Score 1236;	DB 1; Length 666;
Best Local Similarity	39.4%;	Pred. No. 2.50e-96;	
Matches	222;	Conservative 151;	Mismatches 144; Indels 47; Caps 30;
Sequence	666 AA:		
Protein	Location/Qualifiers	1.28	
Peptide	"/note="signal peptide"	29.666	
Protein	"/note="mature protein"		
Protein	Mo9827805-A1.		
PD	02-JUL-1998.		
PR	22-DEC-1997; A00874.		
PR	20-DEC-1996; A0-004275.		
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.		
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;		
PI	WPI: 98-37729/32.		
DR	N-PSDB: V42311.		
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -		
PS	Useful for controlling microbial infestations of plants or mammals		
PS	Claim 1; Page 39-41; 96pp; English.		
CC	The sequence is that of an antimicrobial protein which can		
CC	be used to control microbial infestations in plants and mammalian		
CC	animals.		

RESULT	7	22.3%	Score 966;	DB 1;	Length 605;
ID	W62838 standard; Protein; 605 AA.	Best Local Similarity 34.2%;	Pred. No. 3,32e-72;		
AC	W62838;	Matches 193;	Conservative 168;	Mismatches 164;	Indels 40; Gaps 28;
DT	27-OCT-1998 (first entry)				
DE	Glycine max antimicrobial protein.				
KW	antimicrobial protein; infestation; control.				
OS	Glycine max.				
PN	W09827805-A1.				
PD	02-JUL-1998.				
PF	22-DEC-1997; AU0874.				
PR	20-DEC-1996; AU04275.				
PT	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.				
PI	BOWER NI, Goulter KC, Green JL, Manners JM, Marcus JP;				
DR	WP1: 98-377279/32.				
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -				
PR	useful for controlling microbial infestations of plants or mammals				
PS	Claim 1; Page 63-65; 96pp; English.				
CC	The sequence is that of an antimicrobial protein which can				
CC	be used to control microbial infestations in plants and mammalian				
CC	animals.				
CC	Sequence 605 AA;				
Query Match		22.3%;	Score 966;	DB 1;	Length 605;
Best Local Similarity	34.2%;	Pred. No. 3,32e-72;			
Matches	193;	Conservative	168;	Mismatches	164;
Indels	40;	Gaps	28;		
Db	50 RNQCHARCNU-LKVEKDE--CEGELTPRPAPQHREPERPOQGEKEEDEDEQPRPIPT 106				
Qy	40 RYECRRRCCEMDTGTGQEQEQQCEESCSCSYGKQDQGRHREPERQRYEECCQEQCE- 98				
Db	107 PRPQROEEHQRDEQREEMWPKKEKREKSESEEDDEDEDEDEDEDEDEDEDEDEDEDEDE 166				
Qy	99 ERQPPQCCQRLKRFDE-QQOSRQDQ-CQCHQHQEQREPKQ-QCYRECEKQEN 155				
Db	167 EEDDEDEQRESESESDSELRRHKNKNPLFGSNRFETLKNQYGRIVLQRFNORSPO 226				
Qy	156 PWREERESEA-EETETEAGE-QEOSH-NPFFHRRSFQSRFREDHGNFVYLQRFASRHP 212				
Db	-227 LQNLDRRIIEFNKPTMLLPNADADYLVILNGAITSLVNDDRDYSRLQSGDLR 286				
Qy	213 LRGINERFLSTLENPMTFVLPFHCDKEKITYLVNGSLTFLTHKKESINIVPGYVK 272				
Db	287 VPSGTYVVVDPDNNMLRLITLAPVKNKGRESEFSLSTEAQSYLOGFSRNILEASY 346				
Qy	273 VPASSTYVLANODKKEKIIAIVLRPVPNNPGQFEFPFAGSQRPQSYLRAFSREILPAF 332				
Db	347 DTRKEELINKVLFSNEBQOQGEQQLQSYVLEISKEDIRALSKRAKSSSKRTISSEKPF 406				
Qy	333 NTRSEQLDE-LFGGRQRRR-QQG-Q-GMFKASQEQDIRALSDQA-TSPREK-SGERFAF 386				
Db	407 NLRSRDITYSKLKFEEITPEKPKQRLRDIDFISYDMNEGALLPHFNSKAIVLVIN 466				
Qy	387 NLSQTRIVYSNQNRFFEDACPRPERQLRDINIVYSAQLNQGSLFYHYHYSKAFVILVT 446				
Db	467 EGDANIELVG-L-K---EQQEQ-QQEQPLEV-RK---YR---AELSDQIFVIPIAG 511				
Qy	447 EGNVYAEVWSPHLPROSSYEEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 506				
Db	512 YPVV-VNAT-NL---NF--FAIGINMENNORNFLLASQONVISOIVQELAFPSAQ 564				
Qy	507 FPVYFVASQNNLNMGTGFLYNINIDHQRILFVACKINH-V-RQWDSQAKELAFGVSSR 565				
Db	565 AVERLLK-NQRESYFVDAQPKKEE 588				
Qy	566 LVDEIFNSNPQESYFVSQRQRASE 590				
RESULT	8				
ID	W62835 standard; Protein; 593 AA.				
AC	W62835;				

[illegible]

Query Match	18.1%	Score 784	DB 1	Length 626
Best Local Similarity	36.3%	Pred. No. 4,56e-56		
Matches 159	Conservative 117	Mismatches 129	Indels 33	Gaps 21
Protein 23.626	/label= Mlt-protein			
Modified_site 521.523	/note="N-glycosylation site"			
WT				
FT				
ET				
FN	W09724139-A1			
PD	10-JUL-1997			
PF	23-SEP-1996	U15222		
PR	04-MAR-1996	US-610424		
PR	29-DEC-1995	US-009455		
PA	(UYAR)- UNIV ARKANSAS			
PI	Bannon GA, Burks AW,			
DR	WPI: 97-363453/33			
DR	N-PSDB: T76613			
PT	Peanut allergens Ara hi and Ara hii - used for vaccination and in			
PS	two-site monoclonal antibody based ELISA			
CC	Claim 31; Page 172; 354pp; English.			
CC	This polypeptide comprises major peanut allergen Ara hi (W2149).			
CC	its sequence was deduced from cDNA clone p41b (T76613), isolated			
CC	from peanut seed cDNA using a primer (see T76616) based on an			
CC	isolated Ara hi peptide (see W24206). The sequence shows			
CC	significant homology with the vicilin family of seed storage			
CC	proteins of other legumes. The allergen is recognised by serum			
CC	IgE from a large proportion of individuals with peanut			
CC	hypersensitivity. Ara hi and Ara hii (see W24164) can be used to			
CC	raise monoclonal antibodies which are used in a specific two-site			
CC	Mab ELISA for the detection of Ara hi or Ara hii (claimed). IgE-			
CC	vaccinating Ara hi antigen epitopes (see W24165-87) may be used in			
CC	vacines to protect against allergic reactions to peanut allergens,			
CC	e.g. anaphylactic shock.			
SO	Sequence 626 AA;			
Query Match	18.1%	Score 784	DB 1	Length 626
Best Local Similarity	36.3%	Pred. No. 4,56e-56		
Matches 159	Conservative 117	Mismatches 129	Indels 33	Gaps 21
Protein 23.626	/label= Mlt-protein			
Modified_site 521.523	/note="N-glycosylation site"			
WT				
FT				
ET				
FN	W09724139-A1			
PD	10-JUL-1997			
PF	23-SEP-1996	U15222		
PR	04-MAR-1996	US-610424		
PR	29-DEC-1995	US-009455		
PA	(UYAR)- UNIV ARKANSAS			
PI	Bannon GA, Burks AW,			
DR	WPI: 97-363453/33			
DR	N-PSDB: T76613			
PT	Peanut allergens Ara hi and Ara hii - used for vaccination and in			
PS	two-site monoclonal antibody based ELISA			
CC	Claim 31; Page 172; 354pp; English.			
CC	This polypeptide comprises major peanut allergen Ara hi (W2149).			
CC	its sequence was deduced from cDNA clone p41b (T76613), isolated			
CC	from peanut seed cDNA using a primer (see T76616) based on an			
CC	isolated Ara hi peptide (see W24206). The sequence shows			
CC	significant homology with the vicilin family of seed storage			
CC	proteins of other legumes. The allergen is recognised by serum			
CC	IgE from a large proportion of individuals with peanut			
CC	hypersensitivity. Ara hi and Ara hii (see W24164) can be used to			
CC	raise monoclonal antibodies which are used in a specific two-site			
CC	Mab ELISA for the detection of Ara hi or Ara hii (claimed). IgE-			
CC	vaccinating Ara hi antigen epitopes (see W24165-87) may be used in			
CC	vacines to protect against allergic reactions to peanut allergens,			
CC	e.g. anaphylactic shock.			
SO	Sequence 626 AA;			

